CLAIMS

What is claimed is:

1. 1. A method for comparing a media sample and a media file, comprising:
2. computing a set of sample fingerprints, each sample fingerprint characterizing a
3. particular location within said media sample;
4. obtaining a set of file fingerprints, each file fingerprint characterizing at least one file
5. location within said media file;
6. generating correspondences between said particular locations of said media sample and
7. said file locations of said media file, wherein corresponding locations have
8. equivalent fingerprints; and
9. identifying said media sample and said media file if a plurality of said corresponding
10. locations are substantially linearly related.

1. 2. The method of claim 1 wherein said particular locations within said media sample
2. are computed in dependence on said media sample.

1. 3. The method of claim 1 wherein each sample fingerprint represents one or more
2. features of said media sample near said particular location.

1. 4. The method of claim 1 wherein said sample fingerprints and said file fingerprints
2. have numerical values.

1. 5. The method of claim 1 wherein values of said sample fingerprints specify a
2. method for computing said sample fingerprints.

1. 6. The method of claim 1 wherein said media sample is an audio sample.

1. 7. The method of claim 6 wherein said particular locations are timepoints
2. within said audio sample.

1. 8. The method of claim 7 wherein said timepoints occur at local maxima
2. of spectral Lp norms of said audio sample.

1
2

9. The method of claim 6 wherein said sample fingerprints are computed from a frequency analysis of said audio sample.

1
2
3

10. The method of claim 6 wherein said sample fingerprints are selected from the group consisting of spectral slice fingerprints, LPC coefficients, and cepstral coefficients.

1
2

11. The method of claim 6 wherein said sample fingerprints are computed from a spectrogram of said audio sample.

1
2
3
4

12. The method of claim 11 wherein salient points of said spectrogram comprise time coordinates and frequency coordinates, and wherein said particular locations are computed from said time coordinates, and said fingerprints are computed from said frequency coordinates.

1
2
3
4
5
6

13. The method of claim 12, further comprising linking a plurality of said salient points to an anchor salient point, wherein one of said particular locations is computed from a time coordinate of said anchor salient point, and a corresponding fingerprint is computed from frequency coordinates of at least one of said linked salient points and said anchor point.

1
2

14. The method of claim 13, wherein said linked salient points fall within a target zone.

1
2

15. The method of claim 14, wherein said target zone is defined by a time range.

1
2

16. The method of claim 14, wherein said target zone is defined by a frequency range.

1
2

17. The method of claim 14, wherein said target zone is variable.

1
2
3
4
5

18. The method of claim 13 wherein said corresponding fingerprint is computed from a quotient between two of said frequency coordinates of said linked salient points and said anchor point, whereby said corresponding fingerprint is time-stretch invariant.

1
2
3
4

19. The method of claim 13 wherein said corresponding fingerprint is further computed from at least one time difference between said time coordinate of said anchor point and said time coordinates of said linked salient points.

1
2
3
4
5
6

20. The method of claim 19, wherein said corresponding fingerprint is further computed from a product of one of said time differences and one of said frequency coordinates of said linked salient points and said anchor point, whereby said corresponding fingerprint is time-stretch invariant.

1
2
3
4

21. The method of claim 6 wherein said particular locations and said sample fingerprints are computed from salient points of a multidimensional function of said audio sample, wherein at least one of said dimensions is a time dimension and at least one of said dimensions is a non-time dimension.

1
2

22. The method of claim 21 wherein said particular locations are computed from said time dimensions.

1
2

23. The method of claim 21 wherein said sample fingerprints are computed from at least one of said non-time dimensions.

1
2
3

24. The method of claim 21 wherein said salient points are selected from the group consisting of local maxima, local minima, and zero crossings of said multidimensional function.

1
2
25.	The method of claim 6 wherein said sample fingerprints are time-stretch invariant.

1
2
26.	The method of claim 6 wherein each sample fingerprint is computed from multiple timeslices of said audio sample.

1
2
27.	The method of claim 26 wherein said multiple timeslices are offset by a variable amount of time.

1
2
28.	The method of claim 27 wherein each fingerprint is computed in part from said variable amount.

1
2
29.	The method of claim 1 wherein said identifying step comprises locating a diagonal line within a scatter plot of said corresponding locations.

1
2
30.	The method of claim 29 wherein locating said diagonal line comprises forming differences between said corresponding locations.

1
2
31.	The method of claim 30 wherein locating said diagonal line further comprises sorting said differences.

1
2
32.	The method of claim 30 wherein locating said diagonal line further comprises calculating the peak of a histogram of said differences.

1
2
33.	The method of claim 1 wherein said identifying step comprises computing one of a Hough transform and a Radon transform of said correspondences.

1
2
34.	The method of claim 33 wherein said identifying step further comprises locating a peak of said Hough transform.

1
2
35.	The method of claim 1 wherein said identifying step comprises determining whether said the number of correspondences exceeds a threshold value.

1
36.	The method of claim 1 further comprising:

2   obtaining from a database index additional fingerprints characterizing file
3   locations of additional media files;
4   generating additional correspondences between said particular locations of said
5   media sample and said file locations of said additional media files, wherein
6   corresponding locations have equivalent fingerprints; and
7   identifying media files for which a plurality of said corresponding locations are
8   substantially linearly related.

1   37. The method of claim 36 further comprising selecting a winning media file
2   from said identified media files, wherein said winning media file has a
3   largest plurality of substantially linearly related corresponding locations.

1   38. The method of claim 36 wherein identifying said media files for which a
2   plurality of said corresponding locations are substantially linearly related
3   comprises searching a first subset of said additional media files.

1   39. The method of claim 38 wherein additional media files in said first
2   subset have a higher probability of being identified than additional
3   media files that are not in said first subset.

1   40. The method of claim 39 wherein said probability of being
2   identified is computed in dependence on a recency of previous
3   identification.

1   41. The method of claim 39 wherein said probability of being
2   identified is computed in dependence on a frequency of previous
3   identification.

1   42. The method of claim 38 wherein identifying said media files for which
2   a plurality of said corresponding locations are substantially linearly
3   related further comprises searching a second subset of said additional
4   media files, wherein said second subset is searched if no media files in
5   said first subset are identified.

1
2

43. The method of claim 36, further comprising ranking said additional media files according to a probability of being identified.

1
2

44. The method of claim 43 wherein said probability is computed in part in dependence on a recency of previous identification.

1
2
3

45. The method of claim 44 wherein said probability is computed in part by increasing a recency score of a particular media file when said particular media file is identified.

1
2
3

46. The method of claim 44 wherein said probability is computed in part by decreasing recency scores of said additional media files at regular time intervals.

1
2

47. The method of claim 46 wherein said recency scores are decreased exponentially in time.

1
2
3
4

48. The method of claim 43 wherein identifying said media files for which a plurality of said corresponding locations are substantially linearly related comprises searching said additional media files according to said ranking.

1
2
3
4
5

49. The method of claim 36 wherein identifying said media files for which a plurality of said corresponding locations are substantially linearly related comprises terminating said search at a media file having a number of said substantially linearly related corresponding locations that exceeds a predetermined threshold.

1
2

50. The method of claim 1 wherein said method is implemented in a distributed system.

1
2

51. The method of claim 50 wherein said computing step is performed in a client device, said obtaining, generating, and identifying steps are performed in a

3   central location, and the method further comprises transmitting said sample

4   fingerprints from said client device to said central location.

1   52.   The method of claim 1, further comprising repeating said computing, obtaining,

2   generating, and identifying steps for sequentially growing segments of said media

3   sample.

1   53.   The method of claim 1 wherein said obtaining, generating, and identifying steps

2   are performed at periodic intervals on a rolling buffer storing said computed

3   fingerprints.

1   54.   The method of claim 1, further comprising obtaining said media sample, wherein

2   said computing step and said obtaining step are performed simultaneously.

1   55.   A method for comparing a media sample and a media file, comprising:

2   receiving a set of sample fingerprints, each sample fingerprint characterizing a particular

3   location within said media sample;

4   obtaining a set of file fingerprints, each file fingerprint characterizing at least one file

5   location within said media file;

6   generating correspondences between said particular locations of said media sample and

7   said file locations of said media files, wherein corresponding locations have

8   equivalent fingerprints; and

9   identifying said media sample and said media file if a plurality of said corresponding

10   locations are substantially linearly related.

1   56.   A method for recognizing a media sample, comprising:

2   computing a set of sample fingerprints characterizing a segment of said media sample;

3   storing said fingerprints in a rolling buffer;

4   obtaining a set of matching fingerprints in a database index, each matching fingerprint

5   characterizing at least one media file and matching at least one fingerprint in said

6   rolling buffer;

7   identifying at least one media file having a plurality of matching fingerprints; and

8   removing at least one sample fingerprint from said rolling buffer.

1   57.   The method of claim 56, further comprising repeating said method for additional
2   segments of said media sample.

1   58.   The method of claim 56 wherein said computing, storing, and removing steps are
2   performed in a client device and said locating and identifying steps are performed
3   in a central location, and wherein the method further comprises transmitting said
4   sample fingerprints from said client device to said central location.

1   59.   The method of claim 56 wherein said computing step is performed in a client
2   device and said storing, locating, identifying, and removing steps are performed in
3   a central location, and wherein the method further comprises transmitting said
4   fingerprints from said client device to said central location.

1   60.   A method for characterizing an audio sample, comprising:
2   computing a set of reproducible locations in said audio sample; and
3   computing a set of fingerprints characterizing said reproducible locations in said audio
4   sample.

1   61.   The method of claim 60, wherein said reproducible locations and said fingerprints
2   are computed simultaneously.

1   62.   A program storage device accessible by a computer, tangibly embodying a program of
2   instructions executable by said computer to perform method steps for comparing a
3   media sample and a media file, said method steps comprising:
4   computing a set of sample fingerprints, each sample fingerprint characterizing a
5   particular location within said media sample;
6   obtaining a set of file fingerprints, each file fingerprint characterizing at least one file
7   location within said media file;
8   generating correspondences between said particular locations of said media sample and
9   said file locations of said media file, wherein corresponding locations have
10   equivalent fingerprints; and
11   identifying said media sample and said media file if a plurality of said corresponding
12   locations are substantially linearly related.

63. A system for recognizing a media sample, comprising:

a landmarking and fingerprinting object for computing a set of particular locations within said media sample and a set of sample fingerprints, each sample fingerprint characterizing one of said particular locations ;

a database index containing file locations and corresponding file fingerprints for at least one media file; and

an analysis object for:

locating a set of matching fingerprints in said database index, wherein said matching fingerprints are equivalent to said sample fingerprints;

generating correspondences between said particular locations of said media sample and file locations of said at least one media file, wherein corresponding locations have equivalent fingerprints; and

identifying at least one media file for which a plurality of said corresponding locations are substantially linearly related.

64. A computer-implemented method for creating a database index of at least one audio file in a database, comprising:

computing a set of fingerprints representing features of each audio file, each fingerprint characterizing a particular location within said audio file; and

storing within a memory said fingerprints, said locations, and an identifier of each media file, wherein each corresponding fingerprint, location and identifier is associated in said memory.

65. The method of claim 64, further comprising sorting said database index by fingerprint value.

66. The method of claim 64 wherein said particular locations of each audio file are computed in dependence on said audio file.

67. The method of claim 64 wherein each fingerprint represents at least one feature of said audio file near said particular location.

68. The method of claim 64 wherein said fingerprints are numerical values.

1
2
69. The method of claim 64 wherein values of said fingerprints specify a method for computing said fingerprints.

1
2
70. The method of claim 64 wherein said particular locations are timepoints within said audio file.

1
2
71. The method of claim 70 wherein said timepoints occur at local maxima of spectral Lp norms of said audio file.

1
2
72. The method of claim 64 wherein said fingerprints are computed from a frequency analysis of said audio file.

1
2
73. The method of claim 64 wherein said fingerprints are selected from the group consisting of spectral slice fingerprints, LPC coefficients, and cepstral coefficients.

1
2
74. The method of claim 64 wherein said fingerprints are computed from a spectrogram of said audio file.

1
2
3
4
75. The method of claim 74 wherein salient points of said spectrogram comprise time coordinates and frequency coordinates, and wherein said particular locations are computed from said time coordinates, and said fingerprints are computed from said frequency coordinates.

1
2
3
4
5
6
76. The method of claim 75, further comprising linking a plurality of said salient points to an anchor salient point, wherein one of said particular locations is computed from a time coordinate of said anchor salient point, and a corresponding fingerprint is computed from frequency coordinates of at least one of said linked salient points and said anchor point.

1
2
77. The method of claim 76, wherein said linked salient points fall within a target zone.

78. The method of claim 77, wherein said target zone is defined by a time range.

79. The method of claim 77, wherein said target zone is defined by a frequency range.

80. The method of claim 77, wherein said target zone is variable.

81. The method of claim 76, wherein said corresponding fingerprint is computed from a quotient between two of said frequency coordinates of said linked salient points and said anchor point, whereby said corresponding fingerprint is time-stretch invariant.

82. The method of claim 76, wherein said corresponding fingerprint is further computed from at least one time difference between said time coordinate of said anchor point and said time coordinates of said linked salient points.

83. The method of claim 82, wherein said corresponding fingerprint is further computed from a product of one of said time differences and one of said frequency coordinates of said linked salient points and said anchor point, whereby said corresponding fingerprint is time-stretch invariant.

84. The method of claim 64 wherein said particular locations and said fingerprints are computed from salient points of a multidimensional function of said audio file, wherein at least one of said dimensions is a time dimension and at least one of said dimensions is a non-time dimension.

85. The method of claim 84 wherein said particular locations are computed from said time dimensions.

1
2
86. The method of claim 84 wherein said fingerprints are computed from at least one of said non-time dimensions.

1
2
3
87. The method of claim 84 wherein said salient points are selected from the group consisting of local maxima, local minima, and zero crossings of said multidimensional function.

1
88. The method of claim 64 wherein said fingerprints are time-stretch invariant.

1
2
89. The method of claim 64 wherein each fingerprint is computed from multiple timeslices of said audio file.

1
2
90. The method of claim 89 wherein said multiple timeslices are offset by a variable amount of time.

1
2
91. The method of claim 90 wherein said fingerprints are computed in part from said variable amounts.

1
2
3
4
5
6
7
92. A method for recognizing a media sample, comprising:
for each of a plurality of media files, providing a file representation of said media file;
providing a sample representation of said media sample; and
identifying at least one similar file representation among said file representations, wherein said similar file representation is similar to said sample representation, by searching said file representations, wherein said searching is performed in part in dependence on a probability of identification of said file representations.

1
2
93. The method of claim 92 wherein said at least one similar file representation exceeds a threshold similarity to said sample representation.

1
2
3
4
94. The method of claim 92 wherein said identifying step comprises searching a first subset of said file representations, wherein said first subset contains file representations having a higher probability of identification than file representations that are not in said first subset.

95. The method of claim 94, further comprising searching a second subset of said file representations if said first subset does not comprise said at least one similar file representation.

96. The method of claim 92, further comprising ranking said file representations by said probability of identification, wherein said identifying step comprising searching said file representations in order of said ranking.

97. The method of claim 96, further comprising terminating said search when said at least one similar file representation is identified.

98. The method of claim 92 wherein said probability of identification is computed in part in dependence on a recency of previous identification.

99. The method of claim 98 wherein a recency score of a particular file representation is increased when said particular file representation is identified.

100. The method of claim 98 wherein recency scores of said file representations are decreased at regular time intervals.

101. The method of claim 100 wherein said recency scores are decreased exponentially in time.

102. The method of claim 92 wherein said probability of identification is computed in part in dependence on a frequency of previous recognition.

103. A method for recognizing a media sample, comprising identifying media files for which locations of a substantial plurality of equivalent features of said media files and said media sample are substantially linearly related.

104. A method for comparing an audio sample and an audio file, comprising:
for each of at least one audio file, computing a plurality of file fingerprints representing said audio file;

4   computing a plurality of sample fingerprints representing said audio sample; and

5   identifying said audio sample and said audio file if at least a threshold number of said

6   file fingerprints are equivalent to said sample fingerprints;

7   wherein said sample fingerprints are invariant to time stretching of said audio sample.

1   105. The method of claim 104 wherein said sample fingerprints comprise quotients of

2   frequency components of said audio sample.

1   106. The method of claim 104 wherein said sample fingerprints comprise products of

2   frequency components of said audio sample and time differences between points

3   in said audio sample.